Available online at www.sciencedirect.com
ScienceDirect
Physica B 385–386 (2006) 825–830
PHYSICA B
www.elsevier.com/locate/physb

Neutrons reveal how nature uses structural themes and variation in biological regulation

Jill Trewhella*

School of Molecular and Microbial Biosciences, University of Sydney, Sydney 2006, Australia

Abstract

Healthy cellular function requires tight regulation of a multitude of bio-molecular interactions and processes, often in response to external stimuli. In achieving this regulation, nature uses a number of 'second messengers' that are released inside cells in response to first messengers, such as hormones that bind to the cell surface. Divalent calcium and cyclic nucleotides, like cAMP, are among nature's second messengers that bind to receptor proteins inside cells order to regulate the activities of various targets, including many protein kinases. Kinases are enzymes that catalyze the attachment of phosphate groups to proteins in order to modulate their functions. We have been using neutron contrast variation and small-angle solution scattering to study the interactions of the second messenger receptor proteins and their regulatory targets in order to understand the structural basis for these complex processes that use a number of common structural motifs to accomplish highly regulated function. Our most recent work has focused on the different isoforms of the cAMP-dependent protein kinase and the muscle regulatory complex troponin.

Keywords: Protein structure; Biochemical regulation; Neutron contrast variation

1. Neutrons and structural biology at the frontier

Life, at the molecular level, is propelled by the interactions of genes (DNA) and gene products (proteins and RNA) as they carry out highly regulated and coordinated biological functions. Nature utilizes a variety of messenger ions or small molecules (e.g., Ca^{2+} or cyclic nucleotides) to interact with these biological macromolecules to regulate their functions and transmit signals. The completion of the Human Genome Project, hailed as providing the 'code' for human life along with the means to rapidly decode the DNA sequences of all forms of life, has led to international structural genomics projects aimed at defining the three-dimensional structures of large numbers of proteins. Proteins, the ultimate expression of the DNA code, are responsible for the myriad of functions essential for life including communication, movement, respiration, energy transduction, synthesis, degradation, and protection against infection. Proteins accomplish this diversity of function by drawing upon the chemical properties of 20 amino acid building blocks they are made from. A specific protein is defined by its unique sequence of amino acids, strung together in a polypeptide chain that folds into a three-dimensional scaffold that precisely configures the chemical groups required for each protein to carry out its highly specific biochemistry. The developing protein structure data bases have shown us that there are many fewer unique protein scaffolds than there are different proteins. We have known for some time that nature utilizes a common set of chemical reactions, such as reversible phosphorylation, to control a wide variety of functions. These similar chemistries are accomplished by proteins having similar scaffolds. In circumstances where the same function is to be performed, but perhaps with different timing or concentrations of reagents, nature has evolved isoforms of the same protein, i.e., proteins with highly similar structures but with variations in sequence such that the basic biochemistry accomplished is the same but tuned to the needs for a specific location in a cell or for a tissue type.

Beyond the international genome sequencing and structural genomics projects aimed at identifying and

*Tel.: +61 2 93518782; fax: +61 2 093514726.
E-mail address: jtrewhella@usyd.edu.au.

0921-4526/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physb.2006.05.118

determining the structures of the individual protein components of living systems, is the ultimate goal of a molecular level understanding of life: a comprehensive understanding of how proteins interact within molecular networks to transmit signals and regulate bio-molecular function, or within multi-component molecular machines that carry out specialized functions. Neutron scattering occupies a unique niche in the quest for this next level of understanding of complex biological systems. This niche is a consequence of a fortuitous isotope effect involving one of the smallest and most abundant elements in living systems: hydrogen.

2. Contrast variation

Hydrogen is ubiquitous in biology and plays critical roles in enzyme mechanisms and in stabilizing bio-molecular structures via specialized bonding schemes as well as by influencing solvent properties and interactions. Hydrogen atoms are difficult to see using X-rays because they are scattered by electrons, hence the scattering power of an atom is proportional to its number of electrons or atomic number and there are no isotope effects. In contrast, neutrons interact with atomic nuclei and neutron scattering lengths are a complex function of the nucleus/neutron interaction and they vary randomly with atomic and isotopic number. It is the different neutron scattering properties of the isotopes of hydrogen that are the essential underpinning for most of the interesting applications of neutrons in structural biology, particularly those involving the interpretation of coherent, elastic scattering from biological molecules in terms of their three-dimensional structures. In neutron solution scattering applications, which is the focus of this article, these differences are the basis for neutron contrast variation methods that provide unique information on the shapes and dispositions of components within bio-molecular complexes or assemblies.

The small-angle scattering of neutrons (or X-rays) from a macromolecule in solution falls off rapidly from zero angle at a rate that depends upon the size and shape of the macromolecule. The intensity of this scattering profile depends upon the difference in scattering density between the macromolecule and the solvent, i.e., the 'contrast'. Most of the nuclei in biological molecules, as well as the ^{2}H (deuterium) isotope of hydrogen, have positive neutron scattering lengths with similar amplitudes. The ^{1}H isotope of hydrogen is the exception, having about half the amplitude and a negative scattering length due to a phase change between the incident and scattered neutrons. As the neutron scattering density of a molecule is simply the sum of each of its constituent atom's scattering length divided by the total molecular volume, manipulation of $^{1}H/^{2}H$ ratios can dramatically change the average neutron scattering density of a solvent or a macromolecule.

Biological macromolecules can readily be made in deuterated form using bacterial expression systems grown on deuterated media. Because many biological complexes or assemblies can be reconstituted from their components, it is possible to prepare them with specific components deuterated. By then systematically varying the deuterium content of their solvent, one can vary the contribution to the total scattering of the deuterated and non-deuterated components. From measurements of the total scattering as a function of deuterium in the solvent, one can mathematically extract the scattering profiles of the deuterated and non-deuterated components which provide information on the three-dimensional shape of these components. In addition, one obtains a cross term that contains information on their relative dispositions [1]. This information is inherently low resolution, but when combined with high-resolution information on the components of complexes or assemblies it can provide very powerful constraints to aid in computational modeling of these structures while also providing unique insights into associations and conformational transitions during function. While there are examples of molecular complexes and even very large assemblies that have proven amenable to crystallographic analysis (e.g., the ribosome), it is common for interesting complexes and assemblies to show conformational flexibility and dynamic behaviors that make them resistant to crystallization. In such instances, solution scattering methods, which are well suited for studying very large molecular weight systems, can provide unique structural information. Examples follow that show how neutron contrast variation has provided important insights into second-messenger-mediated regulation of protein kinases and muscle action.

3. Protein kinase regulation as seen by neutrons

Protein kinases are enzymes that catalyze the transfer of a phosphate group from nature's principal carrier of biochemical energy, ATP, to a target protein, thus modulating its function. Protein phosphorylation is the most wide-spread mechanism of biochemical regulation and kinases account for almost 2% of the total human genome [2]. The activities of kinases themselves are regulated by the so-called second messengers inside cells. Examples include the Ca^{2+}/calmodulin-dependent kinases, such as myosin light chain kinase (MLCK), and the cAMP-dependent protein kinase (PKA). In each of these examples, kinase activation is controlled by a pseudosubstrate, inhibitor release mechanism in which the binding of either Ca^{2+} or cAMP to a regulatory protein induces a binding event and/or conformational transition that results in the release of a pseudosubstrate from the active site of the kinase thus making it accessible for substrates to bind and be modified.

Neutron and X-ray solution scattering have been used to good effect to understand structural aspects of the Ca^{2+}-calmodulin mediated pseudo-substrate-release mechanism in MLCK. Calmodulin is the major intra-cellular receptor for Ca^{2+} signals in eukaryotes, responsible for the regulation of a diverse array of proteins, including a large number of kinases. MLCK is a classic example of a kinase

having the common catalytic core of the kinase family that is a two-lobed structure with a catalytic cleft between the lobes where substrates (Mg^{2+} ATP and peptide) bind and the phosphoryl-transfer is accomplished. The pseudosubstrate and calmodulin-binding sequence of MLCK is C-terminal to this catalytic core. In its autoinhibited state, the catalytic cleft is blocked by the pseudosubstrate sequence. Fig. 1 shows a schematic summarizing the results of a series of neutron and X-ray scattering studies that reveal details of how the pseudosubstrate is released upon interaction with Ca^{2+}-calmodulin [2–5]. Isolated calmodulin is a dumbbell-shaped molecule with two globular lobes, each containing a pair of Ca^{2+}-binding sites, connected by an extended solvent exposed helix. When four Ca^{2+} ions bind to calmodulin, it in turn binds to MLCK near the catalytic cleft and collapses about a helical sequence segment of the kinase. Calmodulin then translocates to a position remote from the catalytic cleft such that the kinase's autoinhibitory contacts are disrupted. Substrate binding results in a reorientation of calmodulin such that interactions are formed between its N-terminal sequence and the kinase. To date, it has proven impossible to crystallize any complexes of calmodulin with MLCK, and characterization of these structural transitions during activation has only been possible using solution scattering and neutron contrast variation on complexes of MLCK with deuterated calmodulin.

Neutron scattering studies of the multi-functional PKA have revealed details of its regulatory catalytic subunit interactions and the remarkable structural diversity in the different isoforms of this multi-subunit enzyme achieves, in spite of its having high sequence homology and a common domain organization within its sequences. There are four

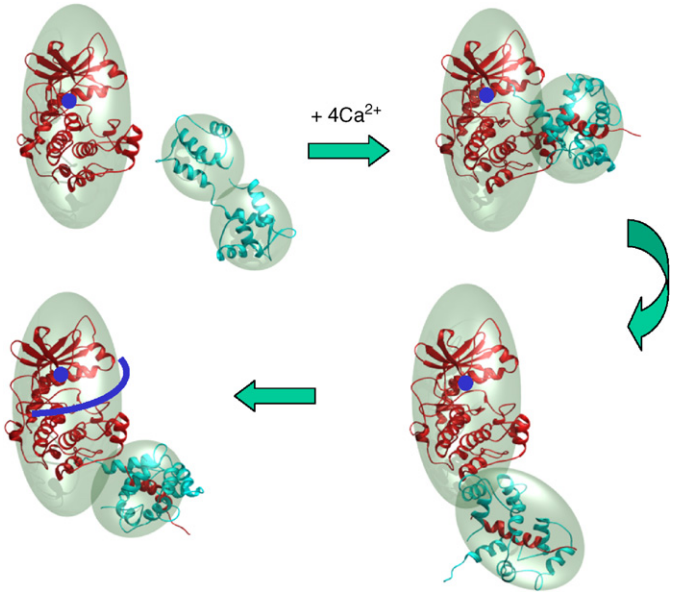

Fig. 1. Schematic of Ca^{2+}-calmodulin (aqua) interacting with MLCK (red). Proteins are shown as backbone representations. Substrates Mg^{2+}ATP and peptide are indicated in blue. Figure adapted from Refs. [5].

known isoforms of PKA (Iα, Iβ, IIα, IIβ) that localize PKA to different parts of a cell via interactions between the R subunit's dimerization/docking domain and membrane-bound receptors known as AKAPs [6]. The inactive PKA holoenzyme is made up of two identical regulatory (R) subunits and two identical catalytic (C) subunits. Starting from its N-terminus, the regulatory subunit has a dimerization/AKAP-docking domain, a linker region which includes a pseudosubstrate/inhibitory sequence, and two tandem cAMP-binding domains. When two cAMP molecules bind to each R subunit, the R-subunit's pseudosubstrate is released by C and substrate binding and modification can proceed. In this mechanism, the R–C interaction is weakened and in some circumstances C dissociates from the R homodimer. The structure of the catalytic subunit of PKA was the first kinase structure to be determined by X-ray crystallography revealing the common catalytic core of the kinase family [7]. There are also crystal structures of the cAMP-binding domains of PKA [8,9], and NMR structures of its N-terminal dimerization domain [10,11], but a flexible linking region between these domains has thus far prohibited high-resolution structural analysis of the intact R subunit or the holoenzyme. The isoforms of PKA are primarily distinguished by differences in their R subunits, specifically in their linker regions and in the sequence connecting the two cAMP-binding domains. We have used neutron and X-ray solution scattering to study the structures of different isoforms of the regulatory dimer of PKA, of the holoenzyme, and of the R–C heterodimer that is missing the R dimerization domain.

Fig. 2 shows a model of the R–C heterodimer interaction from a combination of crystal structures of the components, mutagenesis studies, and neutron contrast variation data from complexes of C with deuterated RIIα [12,13]. The scattering experiments show that in addition to the interaction with the regulatory subunit's pseudosubstrate sequence, there is a secondary interface region. This region is relatively small, as expected for an association poised for dissociation when cAMP and substrate are present. Importantly, both cAMP-binding domains interact with C. A new crystal structure of the C subunit complexed with a single cAMP-binding domain [14] suggests that the details of this interaction need refinement, in particular the orientation of the cAMP-binding domains with respect to C, however the conclusions about the size of the interfaces and the domains involved remain firm. More recent solution scattering studies have revealed that the R–C interaction is different for different isoforms. In the case of the RIα isoform, only one of the cAMP-binding domains interacts with [15].

X-ray solution scattering studies using full-length R subunits revealed details of the regulatory homodimer structures for each PKA isoform [16]. Only the RIα isoform has a well-defined structure for the homodimer (Fig. 3, upper top panel) while the homodimers of the RII isoforms each show evidence of conformational dynamics.

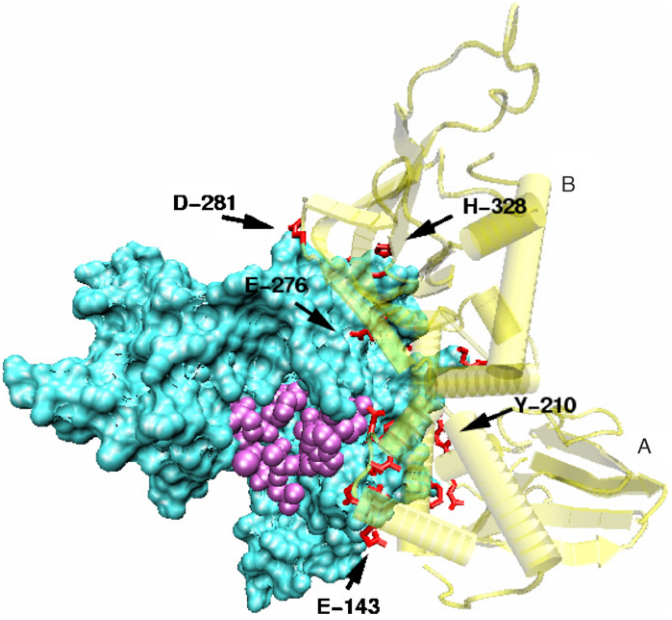

Fig. 2. Model R–C interaction for the type IIα PKA heterodimer. The C subunit is in a space filling representation (aqua with the catalytic cleft highlighted in magenta) and the cAMP-binding domains of R (labelled A and B) in a schematic (yellow) with α-helices and β-strands represented as cylinders and arrows, respectively.

Neutron contrast variation studies of the RIα holoenzyme with deuterated R revealed a conformational change in the homodimer upon C subunit binding (Fig. 3, lower top panel). The cAMP-binding domains are observed to move away from each other and toward the dimerization/AKAP-docking domain [17]. In the RII isoforms, C-subunit binding reduces the conformational flexibility in the R homodimers. Scattering data from three of the four PKA isoforms show how sequence variations in the linker regions of the R subunit facilitate dramatically different spatial arrangements and contacts among the domains and subunits of the holoenzyme (Fig. 3) [18].

The structural diversity seen in the R–C interaction, in the R homodimer, and in the holoenzyme of PKA results in differences in binding surfaces and accessibility that will confer differences in binding affinities, kinetics, and conformational dynamics that are key to understanding the structural basis for the PKA isoforms accomplishing their specific roles in different parts of the cell to which they are targeted via their interactions with the AKAPs. Scattering experiments will continue to contribute to our understanding of this important signaling system.

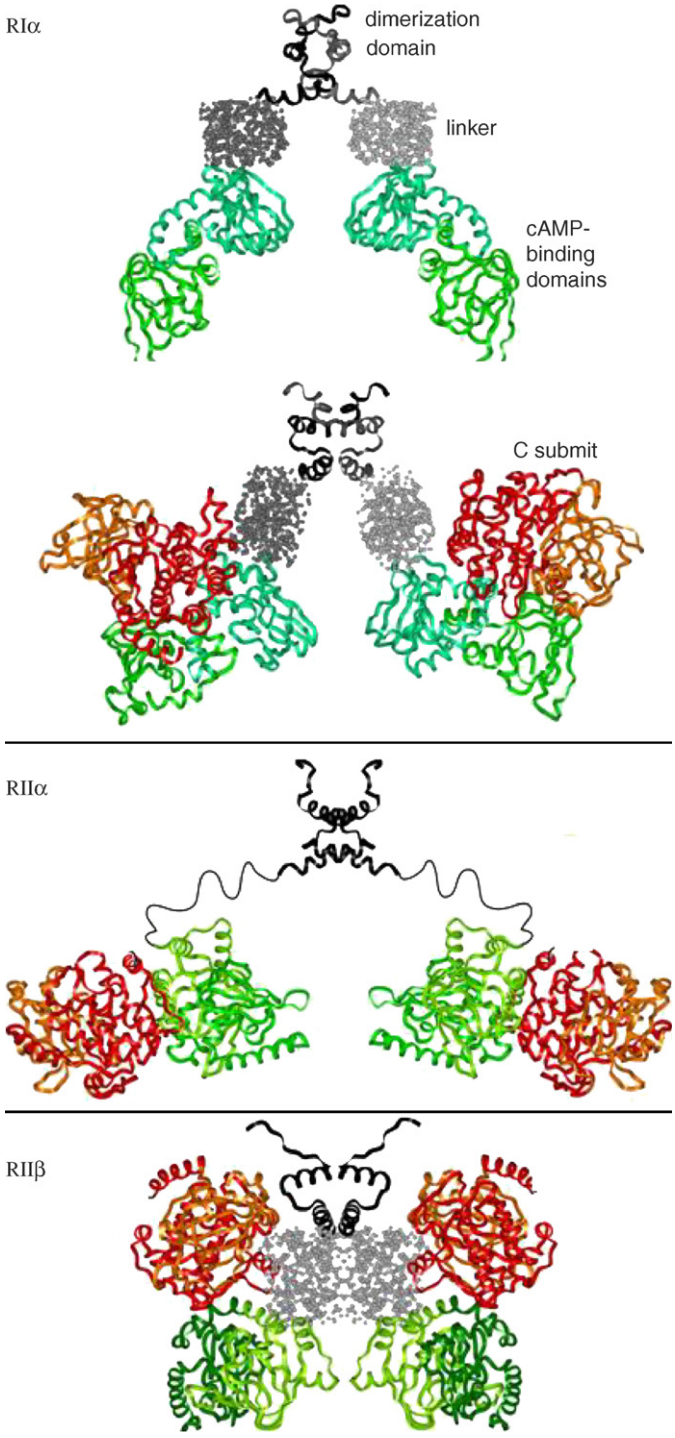

Fig. 3. Structures of three isoforms of PKA. Top: homodimer and holoenzyme structures for RIα, Middle and lower: RIIα and RIIβ holoenzyme structures. Relative scaling of models is approximate. Figure adapted from Refs. [15–8].

4. Neutrons reveal structural details of the regulatory complex troponin in heart and skeletal muscle

Muscle contraction and relaxation occurs when the thick and thin filaments of muscle cells slide passed each other in response to Ca^{2+} signals. The thick filaments are made of highly asymmetric myosin molecules, having long intertwined tails with protruding head groups that form cross bridges between the thick and thin filaments. It is these cross bridges that provide the 'power stroke' for the sliding action via some sort of conformational transition driven by ATP hydrolysis. The thin filaments are made up of actin; globular protein subunits that associate to form a twisted, double-stranded rope-like structure. The regulatory complex troponin sits in grooves on the actin filaments along

with tropomyosin; one troponin and tromomyosin molecule per seven actin monomers. Troponin is the receptor for the Ca^{2+} signals that ultimately regulate the contraction relaxation cycle by controlling the access of the myosin head groups to their actin-binding sites. Tropomyosin aids in transmitting the Ca^{2+} signals received by troponin along the thin filament. Troponin has three subunits: TnC binds Ca^{2+}, TnI inhibits the myosin/actin interaction in the absence of the Ca^{2+} signal, and TnT anchors troponin to the thin filament and aids in transmitting the Ca^{2+} signal. TnC is evolutionarily related to calmodulin, sharing the same overall structure having two globular Ca^{2+}-binding domains, each with two Ca^{2+}-binding sites, connected by a solvent-exposed helix. Unlike calmodulin which has four almost equal affinity Ca^{2+}-binding sites, troponin C has two very high affinity binding sites in its C-terminal lobe that are always occupied in muscle, and lower affinity regulatory sites in its N-terminal lobe that bind Ca^{2+} reversibly during the contraction/relaxation cycle.

Neutron scattering with contrast variation has provided insights into the structure of the troponin complex and differences between its cardiac and skeletal isoforms. Cardiac TnC has one Ca^{2+}-binding site in its N-terminal lobe inactivated, while in the skeletal isoform both of the regulatory N-terminal Ca^{2+}-binding sites are competent. In addition, cardiac TnI (and not the skeletal isoform) has an N-terminal sequence segment containing two serine residues that can be phosphorylated by PKA. This phosphorylation event is believed to impact the equilibrium between the Ca^{2+}-induced conformational changes in the TnC regulatory domain, thereby modulating the Ca^{2+} signal. These isoform differences in TnC and TnI presumably play some role in enabling the involuntary action of cardiac muscle versus the voluntary actions of skeletal muscle. How these differences impact the interactions among the troponin components and with the thin filament has been an important milestone in the quest to understand how this regulatory complex transmits its Ca^{2+} signal.

Small-angle neutron scattering and contrast variation have provided a number of insights into the structural differences between skeletal and cardiac troponin. Fig. 4 shows a model of the binary interaction between skeletal TnC and TnI based on neutron contrast variation studies of the complex formed with deuterated TnC. The model shows skeletal TnC maintains a fully extended, dumbbell shape similar to the crystal structure of the isolated TnC, with TnI twisted about TnC interacting with clefts in each

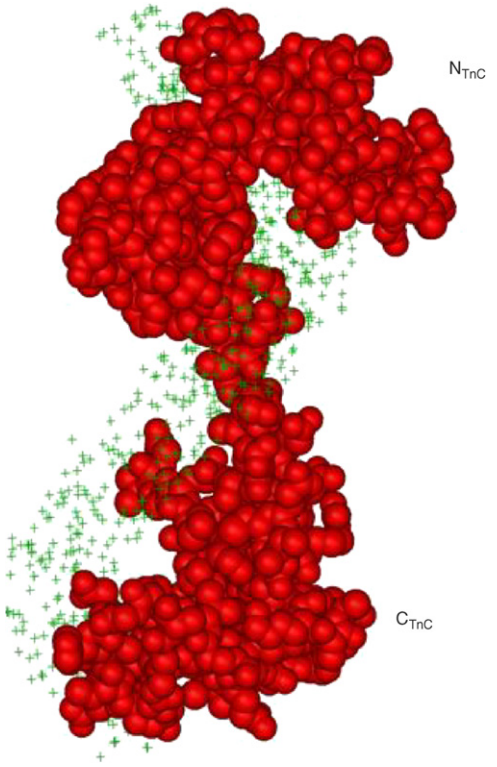

Fig. 4. Skeletal TnC (red, space filling representation based on the crystal structure) complexed with TnI (green dots representing the TnI shape), modeled from neutron contrast variation data and the crystal structure of TnC. Figure adapted from Ref. [20].

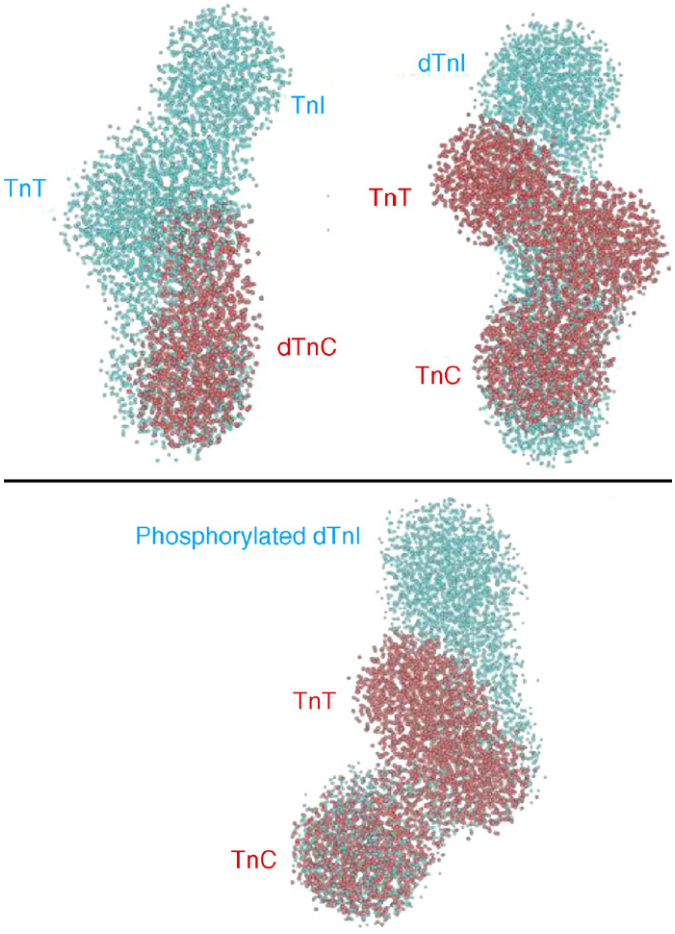

Fig. 5. Shapes of the troponin components derived from neutron contrast variation data with deuterated and non-deuterated components color coded. Top, shows the results of two contrast variation studies; one with deuterated TnC (left, dTnC) and a second with deuterated TnI (right, dTnI). Bottom, shows the conformational changes induced by TnI phosphorylation. Figure adapted from Ref. [24].

Ca^{2+}-binding lobes [19,20]. This result is in stark contrast to what had been previously observed for calmodulin in its interaction with MLCK, and has been subsequently confirmed in neutron scattering studies of the ternary complex [21] as well as a recent crystal structure [22]. These results are leading to the proposal of detailed molecular models for the transmission of the Ca^{2+} signal from TnC to its partners in skeletal muscle thin filament regulation.

Neutron contrast variation studies of the cardiac troponin complex [23,24] revealed a quite different result for this isoform. Fig. 5 summarizes the results of studies in which either TnC or TnI was deuterated in a ternary cardiac troponin complex. In the case of the deuterated TnI experiments, both the phosphorylated and non-phosphorylated forms were studied. The experiment with deuterated TnC shows TnC in a partially collapsed state such that there must be some contacts between the N- and C-terminal Ca^{2+}-binding lobes. Comparison of the shapes derived from samples having the two different deuteration patterns reveals TnT interacting with one lobe of TnC and sitting near the center of a very elongated TnI. Upon phosphorylation, we see TnI undergo a conformational transition in which the part of the structure tightly associated with TnC rotates ($\sim 40^{\circ}$), taking the TnC component with it. These are the first views of the structural effects of phosphorylation of the TnI N-terminal sequence. More work is needed to be able to interpret these changes in terms of atomic models, but they have already provided important clues about the structural differences between cardiac and skeletal troponin that are the basis for their different functional properties.

5. Conclusions

The examples described here demonstrate the kinds of information that can be derived from neutron solution scattering with deuterium labeling and contrast variation. The information content in a solution scattering experiment is inherently low resolution, and there are also limitations that arise from the fact that the molecules are randomly oriented and hence the data are spherically averaged. Interpreting one-dimensional solution scattering profiles reliably in terms of three-dimensional structures is dangerous in the absence of data from other sources. In general, the most valuable small-angle scattering experiments address systems for which high-resolution structural data are available, either on components of complexes, or on the system in one state such that questions can be asked about structural changes induced by a change in state. In these cases, computational modeling, combined where possible with additional structural constraints from mutagenesis, cross-linking, fluorescence data, etc., can be very powerful and lead to detailed molecular models that provide a basis for understanding and predicting molecular function.

Also evident in the examples given here is how nature achieves remarkable structural and functional diversity by evolving variations on a set of common structural themes. Calmodulin and TnC have the same overall three-dimensional fold, but the Ca^{2+}- and target-binding properties in each protein are tuned to fit its specific role. Likewise, the different R-subunit isoforms of PKA have sequence variations that facilitate different spatial domain organizations within this multi-subunit enzyme. These differences impact the R–C interaction, as well as interactions with proteins that are targets for PKA-mediated phosphorylation or anchoring proteins that localize specific isoforms of PKA within the cell. Thus, we see an economy of design in the structural themes and the use of variation in the molecules of life that act in concert to ensure healthy function.

References

[1] M.E. Wall, S.C. Gallagher, J. Trewhella, Ann. Rev. Phys. Chem. 51 (2000) 355.
[2] G. Manning, D.B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, Science 298 (2002) 1912.
[3] J.K. Krueger, G.A. Olah, S.E. Rokop, G. Zhi, J.T. Stull, J. Trewhella, Biochemistry 36 (1997) 6017.
[4] J.K. Krueger, G. Zhi, J.T. Stull, J. Trewhella, Biochemistry 37 (1998) 13997.
[5] J.K. Krueger, S.C. Gallagher, G. Zhi, R. Geguchadze, A. Persechini, J.T. Stull, J. Trewhella, J. Biol. Chem. 276 (2001) 4535.
[6] B.S. Skälhegg, K. Tasken, Front. Biosci. 5 (2000) 678.
[7] D.R. Knighton, J.H. Zheng, L.F. Ten Eyck, V.A. Ashford, N.H. Xuong, S.S. Taylor, J.M. Sowadski, Science 253 (1991) 407.
[8] Y. Su, W.R. Dostmann, F.W. Herberg, K. Durick, N.H. Xuong, L. Ten Eyck, S.S. Taylor, K.I. Varughese, Science 269 (1995) 807.
[9] T.C. Diller, Madhusudan, N.H. Xuong, S.S. Taylor, Structure 9 (2001) 73.
[10] M.G. Newlon, M. Roy, D. Morikis, Z.E. Hausken, V. Coghlan, J.D. Scott, P.A. Jennings, Nat. Struct. Biol. 6 (1999) 222.
[11] P. Banky, M. Roy, M.G. Newlon, D. Morikis, N.M. Haste, S.S. Taylor, P.A. Jennings, J. Mol. Biol. 330 (2003) 1117.
[12] J.K. Zhao, E. Hoye, S. Boylan, D.A. Walsh, J. Trewhella, J. Biol. Chem. 273 (1998) 30448.
[13] C.-S. Tung, D.A. Walsh, J. Trewhella, J. Biol. Chem. 227 (2002) 12423.
[14] C. Kim, N.H. Xuong, S.S. Taylor, Science 307 (2005) 690.
[15] D. Vigil, D.K. Blumenthal, S.S. Taylor, J. Trewhella, J. Biol. Chem. 280 (2005) 35521.
[16] D. Vigil, D.K. Blumenthal, W.T. Heller, S. Brown, J.M. Canaves, S.S. Taylor, J. Trewhella, J. Mol. Biol. 337 (2004) 1183.
[17] W.T. Heller, D. Vigil, S. Brown, D.K. Blumenthal, S.S. Taylor, J. Trewhella, J. Biol. Chem. 279 (2004) 19084.
[18] D. Vigil, D.K. Blumenthal, S.S. Taylor, J. Trewhella, J. Mol. Biol. 357 (2006) 880.
[19] G.A. Olah, S.E. Rokop, C.L.A. Wang, S.L. Blechner, J. Trewhella, Biochemistry 33 (1994) 8233.
[20] G.A. Olah, J. Trewhella, Biochemistry 33 (1994) 12800.
[21] W.A. King, D.B. Stone, P.A. Timmins, T. Narayanan, A.A. von Brasch, P.M. Curmi, J. Mol. Biol. 345 (2005) 797.
[22] M.V. Vinogradova, D.B. Stone, G.G. Malanina, C. Karatzaferi, R. Cooke, R.A. Mendelson, R.J. Fletterick, Proc. Natl. Acad. Sci. USA 102 (2005) 5038.
[23] W.T. Heller, E. Abusamhadneh, N. Finley, P.R. Rosevear, J. Trewhella, Biochemistry 41 (2002) 15654.
[24] W.T. Heller, N. Finley, W.-J. Dong, P. Timmins, H.C. Cheung, P.R. Rosevear, J. Trewhella, Biochemistry 42 (2003) 7790.